# ADDING REALITY TO THE VIRTUAL

Paul J. Metzger

LORAL Advanced Distributed Simulation, Inc.
50 Moulton Street
Cambridge, MA 02138

Distributed interactive simulation has successfully provided an environment for realistic participation in virtual worlds. The technologies behind the generation of these virtual worlds include computer image generation, creation of the polygonal representation of the virtual worlds, and numerous methods for presenting these worlds to the human participants. Humans interact with the virtual world through interface devices such as switches and knobs, keyboards and mice, touch screens and data gloves. The time has come for the seamless integration of these physical, real-world human interface devices with the systems that generate and display the virtual environments. The merging of these two areas will result in virtual world experiences more realistic than any available today.

## 1. INTRODUCTION

Our objective is to introduce more realistic virtual environments and more natural human interfaces to participants in those environments. One way that this can be achieved is by the combination of real world imagery with virtual world environments. We describe here such a system, capable of overlaying images of the real world on top of a virtual world scene, as well as overlaying virtual world imagery onto a real world scene. Applications for both configurations are presented. We conclude with a summary of our research findings to date, and discuss several of the challenges which remain to be confronted with these combined real and virtual systems.

## 2. BACKGROUND: SIMNET

The system described below typically combines "reality" with a synthetic battlefield style of virtual world. The virtual environments we operate in have been generated as a result of technologies and tools developed over the past decade under the SIMulation NETworking, or SIMNET, project.

The SIMNET virtual world features the real time interaction of fully manned simulated vehicles separated by continental distances in real space. Crews interact with realistic controls inside replicated vehicle interior shells, or pods, each equipped with eight to ten vision ports filled with feature-rich computer generated imagery. The crews can interact with other crews in other vehicles by firing weapons and receiving damage, colliding, and exchanging supplies such as fuel and ammunition.

Subsequent programs continue to increase the number of types of distributed simulation entities, which include both manned vehicle simulators and computer-controlled Semi-Automated Forces. For a more thorough discussion of the SIMNET environment, refer to "The SIMNET Architecture for Distributed Interactive Simulation" by Duncan C. Miller [1] or to "The SIMNET Virtual World Architecture" by James Calvin, et. al. [2].



0-7803-1363-1/93 $3.00 © 1993 IEEE

There are two shortcomings of the current SIMNET systems which we've tried to address with the creation of our combined real and virtual system. The first of these shortcomings is that to date, there hasn't been an adequate solution implemented for creating a simulator for training individual foot soldiers in a virtual environment. The arrival of fairly inexpensive and moderate fidelity helmet display systems and tracking devices holds much promise for these applications. However, current display devices make it very difficult to interact with your physical surroundings. You get a very good immersive experience, but you are severely limited in what you can interact with in your immediate physical reality. Even reaching out for a joystick or Spaceball™ device becomes quite a challenge once you've put on the helmet display.

The second shortcoming we've addressed involves recent efforts to link real vehicles on real battle training grounds with networked virtual environments. Instrumenting a tank, for example, and having the real tank show up as a simulated icon in the virtual world can be done (at the 1992 I/ITSEC conference in San Antonio, Texas, LORAL demonstrated this capability). One of the problems remaining with such a system is the challenge of getting the real vehicles to see and interact with vehicles which exist only in the simulated environment.

## 3.0 SYNTHESIZING REAL AND VIRTUAL IMAGERY

The system we present to combine real and virtual worlds addresses both of the above shortcomings. It uses commercially available devices for image processing and display to generate a visual image consisting of real-time computer-generated imagery of the virtual world and live-action video of the real world. This system, developed as part of an internal LORAL virtual reality research project, has two basic modes of operation: the overlay of real-world imagery onto a virtual-world scene, and the overlay of virtual-world imagery onto a real-world scene.

## 3.1 ADDING REALITY TO THE VIRTUAL

A participant enters the cabin of a pod-based simulator. He dons a helmet system equipped with micro-miniature cameras, visual display devices (miniature CRTs or LCDs), and earphones. As the simulation starts, he sees a combination of computer-generated and real-world imagery on the helmet displays. This synthesis of real and virtual imagery allows him to view his real hand (not a disembodied computer-generated model of a hand) interacting with these environments. He is able to see the real-world physical environment of the simulator interior, and can reach out and turn on a switch, watching his real hand perform this task. At the same time, all parts of the simulator designed to provide views into the virtual environment are generating the appropriate computer images.

As Figure 1 shows, the helmet system includes a head tracker, earphones, cameras for live-action video, and visual displays. The CIG produces the virtual world imagery. Between them, and at the heart of this system, is an image processor performing video keying between the real world live-action video and the virtual world imagery.

Video keying is a term which not everyone may be familiar with, but whose use most everyone has had some exposure to. Take the case of a nightly news weather report. The weatherman you see on your TV set is shown standing in front of a computer generated weather map. However, in the TV studio, he's really standing in front of a blue screen. A video keyer is used to overlay the image of the weatherman onto a background image of a computer generated weather map. The virtual reality community may also be familiar with the Mandala system, by Vivid Effects, which incorporates this same form of video keying, commonly referred to as chroma keying.

A composite video signal consists of two parts: the color (or chrominance) signal, and the intensity (or luminance) signal. Video keying can be achieved by either keying on various hues using a chrominance, or chroma, keyer, or by keying against either a black or white background

using a luminance, or luma, keyer. The reason that blue is often chosen as the background color for chroma keyers is due to the fact that this particular shade of blue (often referred to as "chromakey blue") rarely shows up in human skin tones. One could just as easily perform chroma keying against a hot pink or dark green background, if it works better for your application. The tradeoff between chroma and luma keyers is that you can obtain a sharper key using chroma keyers, but luma keyers are generally cheaper (simpler circuitry is involved in luminance keying).

Both forms of video keying work by evaluating an incoming image and removing any part of this foreground image which contains the color you're keying against. These removed segments become "keyholes", allowing you to see through the foreground image to the background image, which has also been input to the video keyer.

In our simulator example, the foreground image captured by the helmet system camera is the real-world physical environment (e.g., the hand turning on the switch). This is analogous to the TV weatherman. The virtual-world CIG imagery provides the background, analogous to the weather map. The live-action video from the helmet system camera (the foreground) is input, along with the real-time CIG imagery (the background), into a video keyer. The keyer processes the images, with the resulting combined output sent as the input signal to the display devices in the helmet system. If the participant were to hold his hand out in front of a "window" in the cabin, with the color of the window being the color used for the video keying, on the helmet displays he would see a live picture of his real hand superimposed on the simulated out-the-window view of the virtual world.

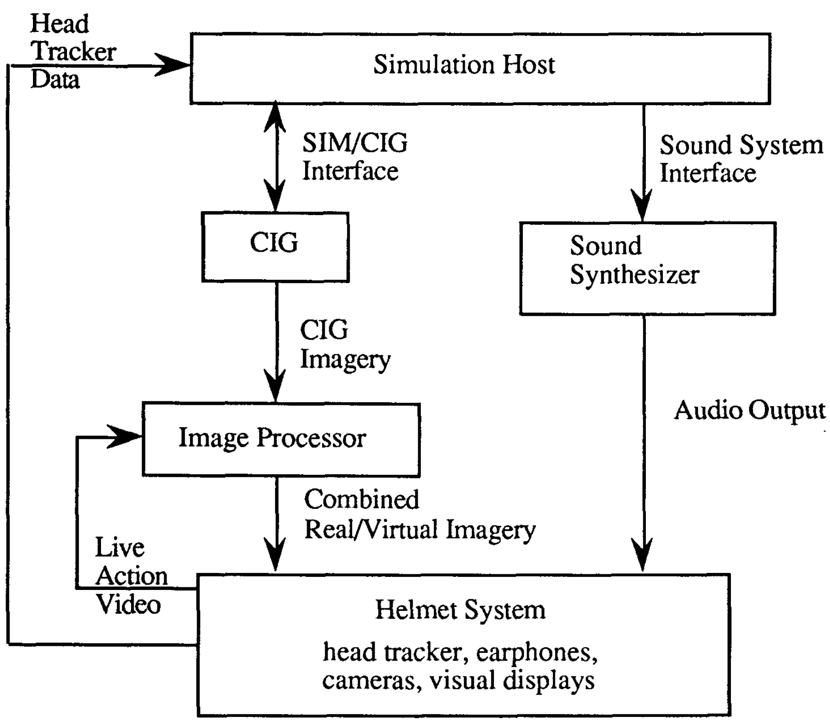


Figure 1: System Overview

A system configuration such as the one just described can be designed to address the problems of creating a simulation environment for individual foot soldiers, for building a very low cost "dome"

style flight simulator, and for making research into virtual environments easier and more productive. This system can enable scientists to view and interact with their physical surroundings while at the same time exploring virtual worlds from within an immersive environment.

In all cases, the central concept is to have real spaces in the physical environment set up with windows into the virtual world which are the colors you want to perform the video keying against. For example, the low cost "dome" flight simulator can be achieved by having the airplane cockpit enclosed within a cube or dome whose walls are painted the color you plan to key against. Instead of needing several expensive projection display devices, many channels of CIG imagery, and lots of floor space to obtain a 360 degree field of regard, you can achieve similar results using a helmet display device, a miniature camera, a video keyer, and one channel of computer imagery (two cameras, keyers, and channels of imagery for stereoscopic views).

## 3.2 AUGMENTING REALITY WITH THE VIRTUAL

The same hardware suite from Figure 1 can be used to insert computer-generated objects from the virtual world into the real world. In this second configuration, the foreground captured by the camera is a CIG virtual world, and the background is the real-world live-action video.

Take the case of trying to inject a virtual helicopter into a real world training ground. The scenario has a simulated helicopter approaching from behind a real hill, then cresting the hill and coming into full view. Our example uses a system incorporating chroma keying, and the color being keyed against is chromakey blue.

This scenario can be achieved by having a virtual world database which models, as accurately as possible, the real world hill. This database contains all of the required polygonal data about the hill, its trees, bushes, etc., but all of the polygons are colored chromakey blue and have no texture maps applied. The simulated sky is also chromakey blue. The helicopter icon is tan, and has a camouflage texture map pattern applied to it.

The helicopter starts in a position fully behind the hill. With the helicopter in this position, and with the viewpoint looking at the hill from the front, the CIG will generate an image of the blue hill against a blue sky — the helicopter is behind the hill, so you can't see it.

A soldier is standing in the real field, looking at the real hill with a pair of "specially equipped binoculars" (cameras mounted on miniature CRTs). The all-blue CIG image is input to the video keyer as the foreground, along with the camera image from the soldiers binoculars as the background. The resultant combined image is one of the real world hill, since the all-blue CIG image becomes one large keyhole, allowing all of the background image to be displayed. This resultant image is fed back to the binoculars as input to the display devices.

The helicopter begins to crest the hill. The CIG generates a mostly blue image, with the exception of the upper parts of a tan camouflage patterned helicopter. The lower parts of the helicopter are masked by the blue hill. The soldier is still viewing the real hill with his specially equipped binoculars. As the virtual helicopter begins to crest the hill, what the soldier sees in his binoculars is a virtual helicopter cresting a real world hill. As the helicopter clears the hill, its complete virtual image comes into full view in the real world scene the soldier is viewing.

## 3.3 SYSTEM COMPONENTS

The helmet system we use consists of the Virtual Research Flight Helmet, an Ascension Technologies Bird position/orientation tracker, and a Toshiba IK-M40A micro-miniature color CCD camera. The camera head is about the size of a lipstick tube, and weighs 0.56 ounces. It's connected via a cable to an external camera control unit, where the electronics are housed. Figure 2

is a photograph of the helmet system. Note that our configuration uses only one camera. For our initial research, we didn't feel it was necessary to generate stereoscopic images. If stereo views are desired, two cameras would need to be mounted on the helmet.

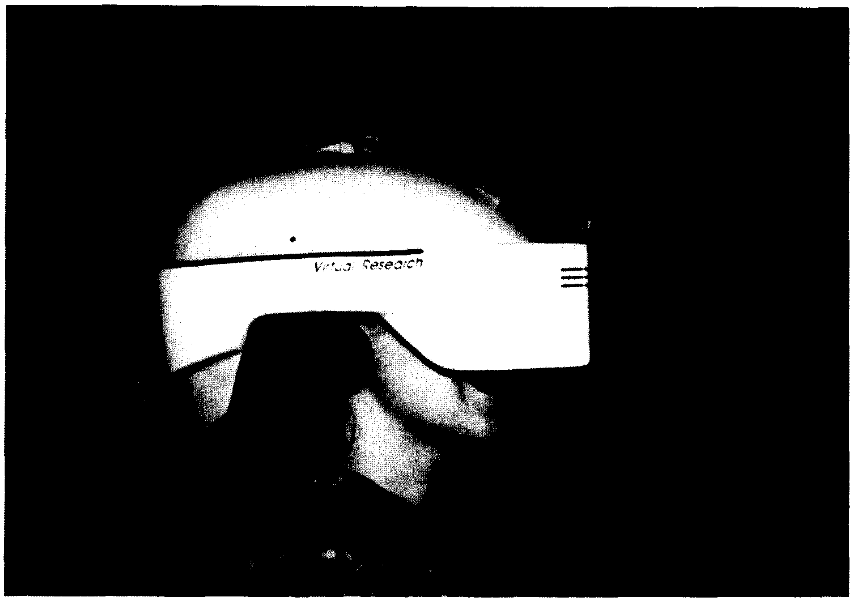

Figure 2: Helmet System

The video keyer we use is a luminance keyer, which is included as a basic component of the NewTek Video Toaster. The Toaster comes with many features which can be used in a combined real and virtual system. If you don't have the capacity to set up parts of your physical environment to be used for video keying, you can use the Toaster to take an input signal, shrink it down to about one fourth of the screen size, and place it anywhere you desire on top of the second input signal (the Toaster is performing video keying internally in this case). This can be used to overlay a small inset of the real world scene on top of the virtual world imagery, without explicitly keying against the physical environment. This is a convenient way to explore immersive virtual worlds while still being able to see and use your Spaceball™, keyboard, or mouse. The Toaster also allows images to be made transparent, so if you wanted to see through a virtual world image to a real world image, you could.

The CIG used to generate the virtual world imagery is a LORAL GT100 series image generator, developed under the SIMNET project. The simulation host is a Motorola MVME147 card, embedded in the GT100 VME bus.

Since the GT100 CIGs produce R-G-B-Sync signals, and the Video Toaster expects an NTSC signal, we run the CIG output through an RGB to NTSC converter prior to input to the Toaster. The Toshiba camera outputs an NTSC signal. The Toaster output is a broadcast quality NTSC signal, which is cabled to the input of the Virtual Research helmet. Figure 3 contains a sample of the types of images which this system produces.

Figure 3: Combining Real and Virtual Imagery

Our configuration also incorporates three dimensional sound localization, although this is an optional feature. We generate sounds using a Yamaha MIDI sequencer and localize sounds using the Focal Point 3D Audio system. The combined audio signals are sent to the headphones equipped with the Virtual Research helmet.

Our system is also connected to an internal SIMNET network. Positional information on the locations of the sound sources is available via the data packets for the vehicles and weapons effects which are present on the network. This can be set up to use either SIMNET, DIS, or IEEE 1278 [3] simulation network protocols.

It should be pointed out that this system is not dependent on any particular type of display device, tracking system, or video keyer. The concepts would work equally well on CRT based helmet displays as with LCD based ones, as well as with boom-mounted displays. Electromagnetic, mechanical, optical, or ultrasonic tracking devices would all work with this system. Most importantly, the type of video keyer you choose can be either a chroma or a luma keyer, based on the requirements of your application and the budget you have available.

## 4.0 FUTURE WORK

While the system we've described in its various configurations appears to have many uses in a variety of applications, there are certain issues which still need further investigation.

One aspect of getting a "good" video key is the importance of lighting conditions. It's important to make sure that what you're keying against is uniformly lit, and does not conflict with colors on your foreground image. With our luminance keyer, we've found that keying against white fluorescent lights works extremely well, due to the high intensity contrast. In a pod-based simulator, a system where the "windows" are back lit lights would work very well with a luminance keyer.

On numerous occasions various references have described how immersive VR systems can create a feeling of vertigo. While we haven't performed any precise human factors studies, our experience with this system leads us to believe that a combined real world / virtual world system can provide enough of an anchor to reality to help minimize the sensation of vertigo. Whenever you start feeling uneasy, instead of having to remove the helmet display or close your eyes, you can look down at your hands and feet, and usually the uneasiness passes.

While lag (or latency) in a helmet display system is undesirable in and of itself, it becomes even more noticeable when you have a low-latency image, such as the image from a live-action video camera, displayed along with the virtual image. A positive aspect of this configuration is that the system becomes a good testbed to evaluate the "How low is enough?" question of reducing latency in a VR system.

In the augmented reality configuration, a major hurdle to be crossed is the registration between the virtual world database and the real world scene. If the virtual world and real world don't match exactly, visual anomalies will occur. The tendency to have discrepancies is much higher in an outdoor setting than it is indoors. Creating a virtual world model of a building interior and having simulated people walking down real world corridors is much easier to achieve without anomaly than having simulated people walking around in an outdoor field.

In addition to the issues of database registration, the augmented reality application of real and virtual systems will need to be fully wireless in order to be used in the majority of applications. The display devices, tracking devices, and cameras will all need the ability to work in a wireless mode.

## 4.0 CONCLUSION

Although there are certain applications in which it is desirable to exclude all reference to reality, there are many applications, particularly in day-to-day interactions with computers and computer simulations, where the goal is to transport you and your immediate reality into the computer generated virtual environment. It will still be quite some time before the line separating the real world and the virtual world disappears completely. However, systems have been created which demonstrate that it is possible to make the distinction between what is real and what is virtual a bit more difficult to determine. The result of such combinations is a greatly enhanced sense of realism to experiences in these virtual environments.

## REFERENCES

[1] Miller, Duncan C. "The SIMNET Architecture for Distributed Interactive Simulation", presented at the Summer Simulation Conference, Baltimore, MD, July 1991.

[2] Calvin, J., Dickens, A., Gaines, B, Metzger, P, Miller, D., and Owen, D. "The SIMNET Virtual World Architecture", presented at the IEEE Virtual Reality Annual International Symposium, Seattle, WA, September, 1993.

[3] Institute of Electrical and Electronics Engineers (IEEE), IEEE P1278 - Standard for Information Technology--Distributed Simulation Application--Process and Data Entity Interchange Formats

For further information on LORAL products and services, write to Mr. J.C. Williams, Vice President of Business Development, 50 Moulton Street, Cambridge, MA, 02138, or via telephone at (617) 873-3577.